Multimedia Systems
Lecture – 20

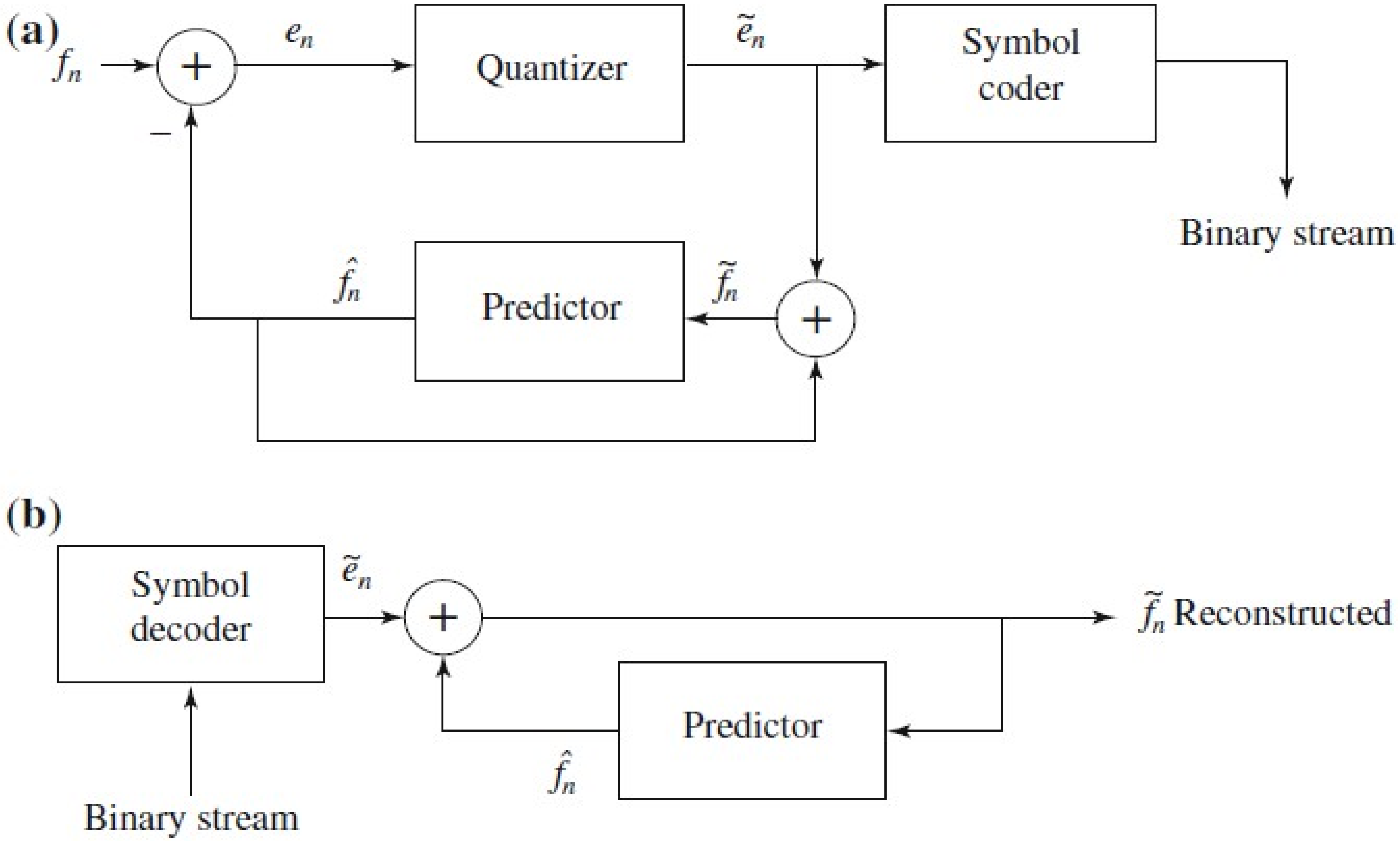
By

Dr. Priyambada Subudhi
Assistant Professor
IIIT Sri City

- Differential Pulse Code Modulation (DPCM) is exactly the same as Predictive Coding, Predictive coding except that it incorporates a quantizer step.
- We shall call the original signal f_n, the predicted signal $\hat{f}_n$, and the quantized, reconstructed signal $\tilde{f}_n$. How DPCM operates is to form the prediction, from an error e_n by subtracting the prediction from the actual signal, then quantize the error to a quantized version, $\tilde{e}_n$.
- The equations that describe DPCM are as follows

$$
\begin{aligned}
\hat{f}_n &= \text{function\_of } (\tilde{f}_{n-1}, \tilde{f}_{n-2}, \tilde{f}_{n-3}, \ldots) \\
e_n &= f_n - \hat{f}_n \\
\tilde{e}_n &= Q[e_n] \\
&\text{transmit codeword}(\tilde{e}_n) \\
&\text{reconstruct: } \tilde{f}_n = \hat{f}_n + \tilde{e}_n
\end{aligned}
$$

Schematic diagram for DPCM: **a** encoder; **b** decoder



- Codewords for quantized error values $\tilde{e}_n$ are produced using entropy coding, such as Huffman coding.
- Notice that the predictor is always based on the reconstructed, quantized version of the signal $\tilde{f}_n$: the reason for this is that then the encoder side is not using any information not available to the decoder side.
- The main effect of the coder–decoder process is to produce reconstructed, quantized signal values $\tilde{f}_n = \hat{f}_n + \tilde{e}_n$
- The distortion is the average squared error $[\sum_{n=1}^{N}(\tilde{f}_n - f_n)^2]/N$
- The predictor makes use of the reconstructed, $\tilde{f}_n$ values not actual signal values f_n—that is, the encoder simulates the decoder in the predictor path. The quantizer can be uniform or nonuniform.

- The prediction value $\hat{f}_n$ is based on however much history the prediction scheme requires: we need to buffer previous values of $\tilde{f}_n$ to form the prediction.
- Notice that the quantization noise, $f_n - \tilde{f}_n$ is equal to the quantization effect on the error term, $e_n - \tilde{e}_n$
- Example: Suppose we adopt a particular predictor as

$$\hat{f}_n = \text{trunc}\left[\left(\tilde{f}_{n-1} + \tilde{f}_{n-2}\right)/2\right]$$

so that $e_n = f_n - \hat{f}_n$ is an integer.

- the particular quantization scheme

$$\tilde{e}_n = Q[e_n] = 16 * \text{trunc}\left[(255 + e_n)/16\right] - 256 + 8$$
$$\tilde{f}_n = \hat{f}_n + \tilde{e}_n$$

- Suppose we wish to code the sequence $f1, f2, f3, f4, f5$ = 130, 150, 140, 200, 230.

- We prepend extra values f_0 = 130 in the datastream that replicate the first value, f_1, and initialize with quantized error $\tilde{e}_1 \equiv 0$, so that we ensure the first reconstructed value is exact: $\tilde{f}_1 = 130$.
- Then subsequent values calculated are as follows

$$\begin{aligned}
\hat{f} &= \boxed{130},\ 130,\ 142,\ 144,\ 167 \\
e &= \boxed{0},\ 20,\ -2,\ 56,\ 63 \\
\tilde{e} &= \boxed{0},\ 24,\ -8,\ 56,\ 56 \\
\tilde{f} &= \boxed{130},\ 154,\ 134,\ 200,\ 223
\end{aligned}$$

Delta Modulation (DM)

- It is a much-simplified version of DPCM often used as a quick analog-to-digital converter.

Uniform-Delta DM

- The idea in DM is to use only a *single* quantized error value, either positive or negative. Such a 1-bit coder thus produces coded output that follows the original signal in a staircase fashion.
- The relevant set of equations is as follows:

$$\hat{f}_{\mathrm{n}} = \tilde{f}_{\mathrm{n}-1}$$

$$e_{\mathrm{n}} = f_{\mathrm{n}} - \hat{f}_{\mathrm{n}} = f_{\mathrm{n}} - \tilde{f}_{\mathrm{n}-1}$$

$$\tilde{e}_{\mathrm{n}} = \begin{cases} +k & \text{if } e_{\mathrm{n}} > 0, \text{ where } k \text{ is a constant} \\ -k & \text{otherwise,} \end{cases}$$

$$\tilde{f}_{\mathrm{n}} = \hat{f}_{n} + \tilde{e}_{\mathrm{n}}$$

- Note that the prediction simply involves a delay.
- Example: Suppose signal values are as follows

f_1	f_2	f_3	f_4
10	11	13	15

- We also define an exact reconstructed value $\tilde{f}_1 = f_1 = 10$.
- Suppose we use a step value $k = 4$. Then we arrive at the following values:

$$\begin{array}{llll} \hat{f}_2 = 10, & e_2 = 11 - 10 = 1, & \tilde{e}_2 = 4, & \tilde{f}_2 = 10 + 4 = 14 \\ \hat{f}_3 = 14, & e_3 = 13 - 14 = -1, & \tilde{e}_3 = -4, & \tilde{f}_3 = 14 - 4 = 10 \\ \hat{f}_4 = 10, & e_4 = 15 - 10 = 5, & \tilde{e}_4 = 4, & \tilde{f}_4 = 10 + 4 = 14 \end{array}$$

- We see that the reconstructed set of values 10, 14, 10, 14 never strays far from the correct set 10, 11, 13, 15.
- It is not difficult to discover that DM copes well with more or less constant signals, but not as well with rapidly changing signals.

Adaptive DM

- However, if the slope of the actual signal curve is high, the staircase approximation cannot keep up.
- A straightforward approach to dealing with a steep curve is to simply change the step size k *adaptively*—that is, in response to the signal's current properties.

Commonly Used Audio Formats

- Digital audio formats emerged with the use and distribution of CD audio discs. These were uncompressed pulse code modulated digital signals in mono and in stereo (*Mono signals are recorded and played back using a single audio channel, while stereo sounds are recorded and played back using two audio channels*).
- However, a number of formats have now become mainstream with the need for streaming, mobile, and surround sound technologies (**Surround sound** is a technique for enriching the fidelity and depth of sound reproduction by using multiple audio channels from speakers that surround the listener (surround channels). Its first application was in movie theaters).

File suffix or logo	Filename	File type	Features
.wav	WAV	Uncompressed PCM coded	Default standard for audio on PCs. WAV files are coded in PCM format.
.au	G.711 μ-law, or ITU μ-law	Uncompressed audio	Universal support for telephone. Packs each 16-bit sample into 8 bits, by using logarithmic table to encode with a 13-bit dynamic range. Encoding and decoding is very fast.
GSM 06.10	Global System for Mobile Communication	Lossy Compressed mobile audio	International standard for cellular telephone technology. Uses linear predictive coding to substantially compress the data. Compression/decompression is slow. Freely available and, thus, widely used
.mp3	MPEG1 Layer3	Compressed audio file format	Uses psychoacoustics for compression Very good bandwidth savings and, hence, used for streaming and Internet downloads.

.ra	Real Audio	Compressed format	Proprietary to Real Audio. Capable of streaming and downloading. Comparable quality to mp3 at high data rates but not so at low data rates
AAC	Advanced Audio Codec MPEG4	Compressed format	Superior quality to .mp3.
.mid	MIDI—Musical Instrument Digital Interface	Descriptive format	MIDI is a language of communication among musical instruments. Description achieved by frequencies, decays, transients, and event lists. Sound has to be synthesized by the instrument.